

## How can you be surprised? The case for volatile expectations

Roberto Casati, Elena Pasquinelli

► **To cite this version:**

Roberto Casati, Elena Pasquinelli. How can you be surprised? The case for volatile expectations. *Phenomenology and Cognitive Sciences.*, 2007, 6, 1-2 (3), pp.171-183. <10.1007/s11097-006-9028-9>. <ijn\_01627163>

**HAL Id: ijn\_01627163**

**[https://jeannicod.ccsd.cnrs.fr/ijn\\_01627163](https://jeannicod.ccsd.cnrs.fr/ijn_01627163)**

Submitted on 5 Nov 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/225736849>

# How can you be surprised? The case for volatile expectations

Article in *Phenomenology and the Cognitive Sciences* · January 2007

DOI: 10.1007/s11097-006-9028-9

---

CITATIONS

10

---

READS

25

2 authors:



**Roberto Casati**

French National Centre for Scientific Research

58 PUBLICATIONS 875 CITATIONS

SEE PROFILE



**Elena Pasquinelli**

Fondation La main à la pâte

13 PUBLICATIONS 36 CITATIONS

SEE PROFILE

Submitted to A. Noë, ed., *Heterophenomenology and Phenomenology*, special issue of *Phenomenology and the Cognitive Sciences*.

## **How can you be surprised? The case for volatile expectations**

Roberto Casati

CNRS-EHESS-ENS, Institut Nicod, Paris and Istituto Universitario di Arti Visive, Venezia

Elena Pasquinelli

EHESS, Institut Nicod, Paris

Submitted version of Oct 13, 2005

### **1. Dennett's plea for surprise**

*“...why do normal perceivers express such surprise when their attention is drawn to facts about the low resolution (and loss of color vision, etc) of their visual peripheries? Surprise is a wonderful dependent variable, and should be used more often in experiments; it is easy to measure and is a telling betrayal of the subject's having expected something else. These expectations are, indeed, an overshooting of the proper expectations of a normally embedded perceiver-agent; people shouldn't have these expectations, but they do. People are shocked, incredulous, dismayed; they often laugh and shriek when I demonstrate the effects to them for the first time. These behavioral responses are themselves data in good standing, and in need of an explanation. They are also, of course, highly reliable signs of their “ideological commitments”.... Surprise is only possible when it upsets belief.” [Dennett, 2001: 982]*

Some concepts and ideas adumbrated in this quote – from a commentary of Dennett's on sensorimotor theories of perception – will be the subject matter of our essay. What is surprise, that it can have a heuristic role? In some experimental settings surprise is explicitly countenanced in measuring subjects' behavioral responses and arguably it can be considered criterial for belief ascription (sections 2-4). The standard explanation of surprise is that it is an emotional reaction to an upset belief; surprise is an epistemologically grounded emotion. Cases in which surprise is reported – in particular, cases of violation of coherence - are discussed (section 5). A distinction between belief and expectation is introduced in order to account for some aspects of surprise; expectations are construed as volatile representations that tie belief to action (section 6). Surprise as invoked by Dennett concerns what he calls 'ideological' commitments; this is a subclass of all the beliefs whose upsetting gives rise to surprise behavior; how do we isolate this class? (section 7) Can lack of surprise provide a complementary heuristic for attributing epistemic states? (section 8). Finally, we shall suggest that surprise provides

Andrea Formica 4/12/06 14:11

Supprimé: 11/27/2006

subjects with some sort of direct phenomenological access to their subpersonal doxastic states (section 9).

## 2. Surprise as an object of empirical investigation

There are a number of domains in which surprise has been investigated as an interesting theoretical item, possibly a natural kind; models have been provided to explain the mechanism of surprise. We can divide the studies on surprise into three main categories: first, studies that are directly devoted to the characterization of surprise; second, studies that are indirectly interested in surprise or that use surprise as an index of the presence of other mental states, cognitive capabilities and processes; third, studies in which the reaction of surprise has been deemed an instrument for evaluating human products, such as theories, in terms of their capacity to go beyond obvious explanations.

The first group of studies directly devoted to the investigation of surprise is mainly represented by the psychological literature on emotions [Darwin, 1872; Bain, 1874; Ribot, 1911; Warren, 1920; Ruckmick, 1921; Woodworth and Schlosberg, 1954; Izard, 1977; Plutchik, 1980] but there are also studies specifically dedicated to surprise [Patterson, 1930; Charlesworth, 1969; Meyer, Niepel, Rudolph and Schützwohl, 1991; Meyer and Niepel, 1994; Niepel, 1994; Stiensmeier-Pelster, 1995; Reisenzein, Meyer and Schützwohl, 1996; Gendolla, 1997; Meyer, Reisenzein and Schützwohl, 1997; Schützwohl, 1998]. Surprise has been characterized as a short-lasting, more or less basic emotion, with no definite positive or negative value. The absence of a value (positive or negative) has led some authors to expunge surprise from the class of emotions, or at least from the basic ones [Mc Dougall, 1923; Ortony, Clore and Collins, 1988]. Nevertheless, as with other emotions, surprise is associated with characteristic physiological and behavioral reactions and subjective sensations that can be measured through subjective reports and scales with objective indexes (such as facial EMG registrations). [Sumitsuji, 2000] for instance seems to prove the existence of a similarity between the facial EMG distribution pattern of the onset of laughter and a state of mild surprise, by electromyography of many facial muscles.

A recent study by [Vanhamme, 2000] dedicated to the measure of surprise in consumers, reveals a lack of correlation between different expressions of surprise: physiological reactions (assessed by frontal EMG for measuring eyebrow-movement, one of the facial expressions most tightly correlated with surprise) show that a subjective report of surprise may easily not be accompanied by significant modifications of the EMG, even if observers notice changes of expression [see Vanhamme, 1999, Vanhamme, 2000 for a synthesis of the methods for measuring surprise]. The study also reveals the possibility of making use of surprise reactions as indicators in the context of adult human behaviors.

The *cause* of surprise is often individuated in some form of unexpected event or unfulfilled anticipation [Darwin, 1872; Plutchik, 1980; Izard, 1991]. Indeed, a number of studies exploit surprise reactions in order to assess the expectations and beliefs possessed by the investigated subjects or to evidence cognitive mechanisms operating in various situations: decision under uncertainty [Coughlan and Connolly, 2001], perception of

Andrea Formica 4/12/06 14:11

Supprimé: 11/27/2006

discrepancy [Whittlesea and Williams, 2001], the bizarreness effect [McDaniel, Einstein, DeLosh, May and Brady, 1995].

This use of reactions of surprise is also apparent in the idea expressed by [Davidson, 1982] that surprise is proof of the mastery of concepts of true and false belief, thus of the concept of belief in general as something that can be false. In Davidson's opinion, it is not possible to be surprised without possessing some beliefs. Conversely, if one possesses some beliefs one is faced with the possibility of being surprised, for something can happen that makes one change one's mind and change one's beliefs. We will discuss this position in section 3.

Analogously, the connection of surprise to expectations grounded in the possession of knowledge has been exploited in the context of infant and adult studies in order to assess reactions to different stimulus presentations and also to individuate the events infants take to be possible and impossible. This aspect too will be further discussed (section 4) in order to show how the reaction of surprise can be exploited to give access to other mental states and cognitive processes.

### 3. Davidson on surprise

Surprise is not only interesting in itself. It is also alleged to be a powerful heuristic for investigating other psychological states. If surprise is an emotion with an epistemological ground, then it is diagnostic of certain underlying doxastic states. In order to be so used, it should be found to be associated systematically with those states. Indeed, surprise has been time and again indicated as a test or a criterion for belief ascription. [Davidson, 1982] has made this point forcefully. He proposes a "surprise test" for errors, whose recognition is in turn a criterion for the ascription of the notion of an extra-mental, objective reality. Someone puts his hand in his pocket and finds a coin. If he is surprised upon finding the coin, then he comes to realize that his previous belief about his pocket was false. Hence, he can be credited with the belief (among others) that there is an objective reality which is independent from (previous) beliefs.

According to Davidson, surprise is the reaction to *frustration of expectation* and for this reason it is a good test for the possession of the concepts of belief, truth and objectivity, since only someone who had a belief, an *expectation*, and realizes themselves to be confronted with a different reality can be surprised.

"Someone who believes there is a dragon in the closet opens the door and sees there is no dragon. He is *surprised*; this is not what he expected. Awareness of the possibility of surprise, the entertainment of expectations – these are essential concomitants of belief."  
[Davidson, 2004, p. 7]

According to Davidson, one cannot be surprised unless one already commands the concept of objective truth. The concepts of truth, of error, objectivity, belief, and the awareness of the possibility of being surprised come in a bundle: if a creature has one of these it has them all.

Andrea Formica 4/12/06 14:11

Supprimé: 11/27/2006

#### 4. Measuring Surprise and Inferring Commitments

In some experimental settings surprise is explicitly considered in measuring subjects' behavioral responses. Two key examples concern the habituation paradigm in the study of infant cognition and in the study of some perceptual illusions.

Infant studies exploit familiarization to a repeated stimulus in order to assess the presence of a reaction toward a different stimulus. A first phase of *habituation* provides the infant with a repeated stimulus. The fact that at a certain moment the attention to the stimulus drops is assumed to indicate that the stimulus has become *familiar*. Once the infant has ceased to pay attention to the stimulus a new stimulus is presented which is different from the preceding one. If the infant recovers attention in the new stimulus then it is possible to assume that the new stimulus has been recognized as different. Arguably, attention recovery can be considered a type of surprise.

Some criticisms have been addressed to the habituation paradigm, e.g. concerns about the problem of infants' preference for familiar stimulations rather than for new ones [Cohen, 2001]. It seems for instance important to take into account the level of complexity of the stimulus: infants would be less and less attracted by a simple stimulus, but they would gain more and more interest in repeating the experience of a complex one. [Spelke, 1990, 1991] and [Baillargeon, 1994] have proposed a modified version of the habituation paradigm which is based on the notion of *violation of expectation*. The increase in attention is intended to signal the perception of an event as impossible. The violation of expectation paradigm is intended to test the infant's knowledge about objects and their properties, such as physical properties. Subsequent to a familiarization or habituation phase, the infant is presented with consistent and inconsistent conditions relative to the situation which has become familiar. It is assumed that the new inconsistent stimuli are considered to be impossible by the infant in that they violate the expectations based upon the beliefs developed (or activated) during the familiarization phase. The impossible event is thus regarded by the infant as novel or surprising, hence more suitable for capturing attention than a consistent event, interpreted as possible, familiar or expected. Infants look reliably longer at inconsistent rather than at consistent stimuli. On the basis of looking time plotted against consistent and inconsistent situations, the authors have unearthed properties such as continuity and impenetrability that characterize the infant's concept of an object [Baillargeon, 1994; Spelke, 1991].

#### 5. Coherence and surprise

Our experience is normally coherent; and we normally expect percepts to cohere. However, not all incoherent or ambiguous experiences would strike us as surprising. In a classical study presented in [Rock and Victor, 1964] and reconstructed by [Heller, Calcaterra, Green and Brown, 1999], the observers are at the same time touching a square object and looking at it through the interposition of minifying lenses that reduces its visual dimensions (the subjects are unaware of this modification, and they assume they are looking at and touching one and the same object). We know that the perceptual

Andrea Formica 4/12/06 14:11

Supprimé: 11/27/2006

system appreciates the difference between the two partial percepts, since the judgments given in the purely visual and in the purely tactile situations are different. How will the perceptual system react in the condition?

A first possibility is that the two partial percepts are not modified. The perceptual system is able to correctly assess the discrepancy between the percepts. The final percept then explicitly contains two discrepant elements and the subject experiences an explicit perceptual conflict. If the conflict is consciously accessed, the subject might report an object as feeling a certain way while looking another.

This doesn't seem to be a common situation. But it is an interesting possibility since the perceptual system can immediately detect the presence of an explicit conflict. Proprioceptive illusions of impossible movements and position could be such a case. [Craske, 1977] describes the effects produced by the vibration of biceps and triceps tendons when the related muscles are stretched against contraction. Some subjects reported a sensation that the arm was in two different places at the same time. The impossibility is related, in this case, not to some form of specific prior knowledge about the object, but to the general fact that there is an intrinsic ambiguity in the final percept. Subjects know that the percept must be wrong because they know that an object cannot be ambiguously located or characterized. When the subjects who had reported unambiguous sensations were newly vibrated and asked to move the limb beyond the point that they had previously reported as the limit of extension or flexion, all of them reported the sensation that the arm was moving beyond the limits of flexion or extension, that is, they reported various degrees of hyperextension and hyperflexion. This sensation was described as follows: "the arm is being broken", "it is being bent backwards, it cannot be where it feels". This situation provokes a reaction of surprise, and is described as impossible by the subject. The subjects also displayed the signs that normally accompany pain, such as writhing, sweating and gasping, even though no pain was actually involved.

A second possible outcome is that the two partial percepts are suitably modified by the cognitive system in order to enter the final unit without giving rise to an explicit conflict. The conflict is solved in favour of a non-ambiguous unity; the competition must have a winner. This solution seems to be more frequent than the one that leads to the experience of explicit conflicts. It seems that the perceptual system has a propensity to compose its experiences into a coherent form, avoiding the possibility that one perceived object can be at the same time, for instance, large for touch and small for vision [Heller, Calcaterra, Green and Brown, 1999].

Finally, the discrepant stimuli can be separated and constitute two distinct percepts, thus avoiding the problem of conflict.

The fact that explicit conflicts are rarely perceived as such, and discrepancies between perceptual or cognitive states are mostly solved into coherent units, indicates that coherence might represent a positive value to be preserved for perceptual and cognitive processes in general. Indeed, perceptual conflicts seem then to have a negative value for programming actions, since they paralyze our decisions and diminish our capacity to behave in the world.

The rarity of explicit experienced conflicts and the reactions in the face of ambiguous stimuli could thus be attributed to a negative value of ambiguity for adaptive behaviours.

Andrea Formica 4/12/06 14:11

Supprimé: 11/27/2006

In this sense the preservation of perceptual coherence represents an adaptive value, and the surprise at the loss of coherence is a valuable signal.

It seems to be possible to extend the considerations about the reactions of the organism to synchronic violations of coherence (when two or more actual stimuli are inconsistent) to the occurrence of diachronic violations of coherence (when inconsistency stands between actual stimuli and past beliefs or knowledge). An experiment performed by [Bruner, 1949] with normal and trick playing cards (i.e. a black three of hearts or a red two of spades) shows that the recognition threshold for the incongruous playing cards is significantly higher than the one for normal cards; four reactions to incongruity are described: either the form or the colour dominates and the subject perceives a normal card, i.e. a normal, red three of hearts instead of a black one, or else a black three of spades; or a compromise object is perceived which composes the conflict, i.e. a greyish three of hearts. Less frequently, the perception of incongruity can also produce disruption, in that the subject cannot solve the recognition task. Finally the incongruity can be recognized. In this case, the recognition of the incongruity is accompanied by a *sense of wrongness* which is analogous to the reaction of surprise and the sense of impossibility which is expressed on the occasion of illusions of impossible movement: the subject suddenly or gradually begins to feel that there is something wrong with the stimulus; this sensation can turn into disruption or give rise to recognition of the incongruity.

Thus it seems that people are not easily disposed to revise their knowledge and past beliefs, at least when well-established expectancies fail confirmation. Nevertheless, surprise only arises when a sense of wrongness or impossibility is elicited by the presence of an explicit conflict or by recognition of the incongruity. This occurrence is not mandatory because different mechanisms can operate that re-establish the coherence of the experience, both at a diachronic and at a synchronic level.

## 6. Surprise, belief and expectation

Davidson characterizes surprise as a reaction to the frustration of expectation. But what exactly is required to model this reaction? What is an expectation, that it can be frustrated? Consider, first, some standard situations in which surprise is (or is not) manifested. Surely there are many things I was not expecting but that do not surprise me at all. Most of what I did today was totally unexpected yesterday, possibly even five minutes ago; but I would not say that it surprised me. I see people walking in the street whose existence I ignore before seeing them. This observation can be turned into a general view about surprise:

(1 Strong Expectation View) Surprise only arises when a *genuine* expectation is violated.

A genuine or strong expectation is to be contrasted with the lack of an expectation. That is: If I do *not* expect that p, the occurrence of p does *not* surprise me. In order to be surprised by p, I need to genuinely *expect* that not-p. I expect to find a soda in the fridge,

Andrea Formica 4/12/06 14:11

Supprimé: 11/27/2006



open the fridge, and find no soda; the Strong Expectation view predicts that I'll be surprised. But it predicts that no surprise will arise when I open the fridge and find no soda if I had no expectation about finding a soda.

However, consider what happens if I have *no* expectation of finding a soda (as contrasted with an expectation of *not* finding a soda), open the fridge, and *do* find a soda. Arguably I'll be surprised by the finding. Or if I travel to a foreign city and meet there, unexpectedly, a schoolmate from my home town. Arguably I did not have the genuine expectation not to find him there, although I did not have the expectation of meeting him there.

The Strong Expectation View can hence be contrasted with:

(2 Weak Expectation View) Surprise arises when an expectation is violated in the sense that a situation occurs in which it turns out that we should have expected something.

That is: If I do not expect that p, the occurrence of p surprises me.

There is a logical link between (1) and (2). If I expect that not-p, then I do not expect that p. A genuine expectation is accompanied by a lack of expectation with polar inverted content. However, on the Strong Expectation View, only the former (the expectation that not-p) is responsible for the surprise.

Something has to be said about expectations. First of all, expectations should be states that are not obviously manipulable at will. If I tell myself, out of the blue, that there is a soda in the fridge, this does not count as an expectation that there is soda in the fridge (and this is what accounts for the fact that I am not surprised if I do not find one). So, expectations look like grounded beliefs, beliefs for which one has some justification available.

Second, however, expectations are not just beliefs. Rather, we may state, beliefs *generate* expectations. An expectation is a belief-grounded representation *that is poised for action*. In this sense, expectations have a content that they inherit from their generating belief. And they have satisfaction conditions that are observed during a particular course of action. However, they are volatile and disappear as soon as the relevant action is accomplished. The input for expectations should be a belief and an action plan in the course of accomplishment.

Third, the distinction between expectation and belief can be deepened by showing that it is orthogonal to the distinction between overt and covert epistemic states. Expectations as well as beliefs can be overt (I tell myself that there are zebras in the savannah), or covert (I have a disposition – a tacit belief whose content is that there are zebras in the savannah). The need for the double distinction is as follows. One can tell oneself that there is a soda in the fridge – and hence have an overt belief that there is a soda in the fridge – but this may not generate an expectation, *if one is not contemplating an action* involving the relevant fridge. On the other hand, one can reach in the fridge during an absent-minded trip to the kitchen, only to find that there is no soda left; one had a covert expectation that there was soda in the fridge, an expectation that did not get

Andrea Formica 4/12/06 14:11

Supprimé: 11/27/2006

verbalized, and whose mismatch with the perceptual judgment that there is no soda in the fridge generates surprise.<sup>1</sup>

To sum up, the rationale for positing expectations as volatile representations is as follows. Surprise requires that accessed representations are present to be matched against the content of an occurrent episode of perception (a perceptual judgment); to that purpose an occurrent belief not sufficient,<sup>2</sup> insofar as it may not be related to the situation in which the perceptual episode takes place. The belief gets connected to the situation when an action involving the objects of the belief is planned or undertaken.<sup>3</sup> The volatile representation, the expectation, is posited to fulfil the role of bringing the belief to action.

## 7. Ideological commitments

Given the characterization of expectation as a volatile representation, on both the Strong and the Weak view surprise arises at the joint of perception and action. This gives rise to a question. I sit on a chair, without *doing* anything. All of a sudden, a gremlin materializes in front of me. I am surprised. Does this violate an expectation? And in the Weak or in the Strong sense?

The belief called into question by the appearance of the gremlin could be of a somewhat general type: I believe that nothing materializes out of the blue.<sup>4</sup> In general, if we have no reason to ascribe the less general of any two beliefs that get violated in a surprise-inducing situation, then we ought to ascribe the more general one. I am surprised not because I do not believe that *Jok*, *this* particular gremlin, or even *a* gremlin can materialize out of the blue, but because I believe that *nothing* can materialize out of the blue.

But this is the content of the underlying belief. Did I *expect* that nothing materializes out of the blue? Surely I do not need to manifest this belief (as we have seen, there can be overt beliefs that generate no expectation), but what can be the criteria for ascribing the expectation? The present account postulates that there is an expectation-generating mechanism, whose inputs are a belief and an action plan. But – or so one may reason – I

---

<sup>1</sup> Barry C. Smith (personal communication) thinks that there may be no room for the distinction between belief and expectation along the lines sketched here. A belief is something that were you to act on it you could use it. The presence/absence of the link to action could be contingent upon the content of the belief. I may have beliefs about events that happen outside the light cone, or about subatomic particles, and these would never be used in an action. However he concedes that one can accept that there are volatile representations when a belief is put into action. If I think that John's bathwater is hot, this does not arouse somatic changes in me. If I think that *this* bathwater is hot, this does arouse some somatic changes. We suggest that these are traces of expectations in the sense here indicated.

<sup>2</sup> Neither is it necessary, insofar as it need not be the case that one explicitly contemplates a certain situation.

<sup>3</sup> Some speculations could be made about the cognitive mechanics of volatile expectations. Hypothetically, a temporary structure such as an object- or an event-file [Kahneman et. al. 1992] is pre-activated when the belief is accessed to control or motivate action.

<sup>4</sup> I may need to revise my belief in the light of this particular materialization, but the issue of belief revision is not the same as the issue of surprise.

was not doing anything; there was no action. Hence there is no expectation that could have been generated relative to that absent action.

We can see what the problem is with the opposition between Weak and Strong. If we follow Strong, there is the risk of expectation inflation – a version of the Frame Problem. If we follow Weak then there is no expectation, but then it is hard to see the justification link between not expecting that p and surprise at p.

We think that we should accept Strong, run the risk of expectation inflation, and judge on a case-by-case basis whether it is reasonable to credit the subject with an expectation that has a given content. The bet is that the inflation will be modest, for quite specific reasons. In the case in question, we hypothesize that a volatile expectation with a very *general* content (“nothing materializes out of the blue”) is generated; hence we are prone to accept that many other such expectations are constantly generated. The generality of the content protects the expectations from inflation. We suggest that these constantly generated expectations are of the type that [Dennett, 2001] calls ‘ideological’ – these can be naïve physical expectations (such as those documented by [Spelke, 1990, 1991]) or naïve psychological expectations; or naïve ontological expectations; or metacognitive expectations such as that we are always aware of all details in a scene (the ones discussed by Dennett). However, this does not require us to think that all possible ideological expectations are generated at any given moment. It may be a highly contextual matter what ideological expectations are generated in a given situation. If we are dealing with material objects (say, when navigating our path over rocky ground) then the relevant expectations are activated, and no situation-irrelevant expectation is activated.

That precisely these ideological expectations are activated is consistent with the fact that expectation is an action-related type of representation. The content of the concerned volatile representations is, supposedly, highly action-sensitive. For instance, naïve physical expectations are relevant in object-involving actions, naïve psychological expectations are relevant in actions involving other people, and metacognitive expectations are relevant in activities such as scanning a scene and exercising conscious attention (as in the cases presented by [Dennett, 2001] or in the experiments on change blindness). When expectations that are relevant for a certain context are violated and inconsistency occurs, the actions that expectations were relevant for can be negatively affected, as the discussion about diachronic and synchronic violations of coherence suggested (section 5).

On the other hand, that not too many expectations, especially not too many ‘anecdotal’ expectations are generated, should be apparent from the fact that we are not, in the norm, overwhelmed by surprise, in spite of the fact that most of what happens to us was not and could not have been expected, without thereby being surprising. At the same time, we are likely to be surprised unerringly by each violation of a high-level expectation.

Andrea Formica 4/12/06 14:11

Supprimé: 11/27/2006

## 8. Negative heuristics: lack of surprise

Alva Noë has suggested that *lack* of surprise can be given equal heuristic significance as that given to surprise. Criticizing Dennett, he states that “for example, the fact that we aren’t surprised by the need to move around and actively look shows that we are not committed to a conception of our own experience as capturing all the detail in a snapshot.” Let us recast the terms of the question on the specific example, where p is the thesis that our experience captures all details in a snapshot, and q is the thesis that we need to move around and actively look.

Dennett’s claim is that:

(1) If you are surprised that not-p, then you are committed to p.

Noë’s claim, on the other hand, is that:

(2) If you are not surprised that q, then you are not committed to p (where p and q have no immediate logical link, but the link can be worked out conceptually).

Now, the problem is that for Dennett:

(3) You are surprised that not-p, hence

(4) You are committed to p.

Whereas for Noë:

(5) You are not surprised by q, hence

(6) You are not committed to p.

Dennett’s conclusion and Noë’s conclusion are at odds. Given we can accept that in the situations under study the factual antecedents (3) and (5) are both true, for subjects are surprised in one case and not in the other, then either conditional principle (1) or (2) (or both) ought to be rejected. We lean towards rejecting Noë’s conditional on grounds of simplicity. Lack of surprise is ubiquitous and it is hard to imagine a robust criterial link with the alleged lack of commitment. I am not surprised that the chair is not talking but this does not show that I am not committed to the belief that chairs do not talk. In a sense, lack of surprise (or epistemic indifference) is not discriminating enough: it cannot ‘see’ the relevant commitment.

## 9. Surprise provides a phenomenology of the subpersonal

The punchline of this final section is that surprise is one peculiar way we possess of accessing some aspects of our standardly unaccessed subpersonal world.

We can distinguish here between third-person access and first-person access, and in the case of first-person access between inferential and direct access. Knowledge gathered in psychological experiments is an example of third-person access to the subpersonal; the understanding of how line drawings work is an example of first-person inferential access, and surprise an example of first-person direct access. The impression that surprise’s being a first-person access to the subpersonal is paradoxical (if it is subpersonal, how can I access it personally?) can be dispelled.

To start from the middle: the case of line drawings provides an example of how we can inferentially access some aspects of our inner life that are usually hidden from knowledge. A line drawing, representing the broad outline of an object, allows one to

Andrea Formica 4/12/06 14:11

Supprimé: 11/27/2006

recognize the object in spite of its schematic character. In particular, the line drawing cannot be taken simply to represent lines in the world – there are no lines out there; nor can it be taken to represent light discontinuities in the scene – there are too many such discontinuities, and a drawing registering all of them will be just unreadable. A hypothesis by Cavanagh (2001) is that adequately composed line drawings are triggers of recognition because there is a level of visual representation in which objects *are* represented as line drawings. When you look at a line drawing you are accessing that particular level of visual information processing; and by recognizing the object of the drawing you can be made to understand the existence of that level. Indeed, it is interesting that the semantic function of line drawings has been known and relied upon for centuries; the absence of a proper theoretical framework just made the fact irrelevant for psychological theorizing. Artists contributed to psychology without realizing it. The access to the subpersonal level of representation is in this case *indirect*; it depends upon an inference whose appropriate (theoretical) premises may or may not be available to the subject.

When you are surprised, the case can be made that this access to some subpersonal functioning is direct. The system uses surprise to warn about some inconsistency in its internal states. In the case of violated expectations, the added value for the system is that surprise signals that more attention should be allocated to the situation. In the case of illusions/conflicts one is aware of, the value comes from the fact that the system is made aware of information that ought to be reassessed. In both cases you are warned that something about yourself ought to be reassessed, and you are presented with the element that ought to be reassessed.

The point of contact between personal access and subpersonal states is, in those cases, only as wide as a point. But it is an interesting ingredient of one's phenomenology.

## 10. Conclusions

There are three dimensions to surprise intended as the result of acknowledging some sort of mismatch in epistemic states. First, a dimension of the explicit vs. implicit character of what gets violated in the mismatch (dispositional vs. occurrent beliefs, as in the distinction between believing without ever being asked that there are zebras in the savannah, and believing it when asked). Second, a dimension of the functional role of what gets violated (belief vs. expectation, the latter intended as volatile representation poised for action, as in the distinction between believing that there is soda in my fridge when I am a thousand miles away and believing it when I am opening the fridge – in the latter case generating an expectation that is absent in the former). Third, a dimension of type of content of what gets violated (anecdotal content vs. ideological content, as in the distinction between expecting that there is a soda in the fridge and expecting that material objects are impenetrable). The case made by Dennett is of this latter type: consider what happens when one sees the change blindness or inattentional blindness stimuli a second time over, after having fallen prey to the illusion on the first pass. One notices that one has been fooled, and one expresses surprise. According to Dennett, surprise reveals a very general underlying belief that one can see all the details of a scene at a glance. The latter is a sort of metacognitive belief, and is not related to the anecdotal content of the situation presented to the subject.

Andrea Formica 4/12/06 14:11

Supprimé: 11/27/2006

Surprise is an occurrent emotional state; the emotion would arise from a comparison between an experienced situation and a pre-existing, conflicting representation; it is grounded in a judgment of mismatch. The problem with surprise is that if it depends on a violation of the pre-existing representation, then it can only arise when the latter is active – be it recalled from long term memory or created on the fly. This indicates two things: First, that it is a volatile expectation for action that is violated, not an occurrent belief, as it is in the experience that surprise is produced. (I can explicitly believe that there is a soda in the fridge at my place, but if when holding this belief I open the fridge at *your* place and do not find a soda I am not likely to be surprised). Second, that dispositional beliefs can give rise to volatile expectations even in the absence of a corresponding occurrent belief (when I have the surprising experience as of a person walking through a brick wall, I may explicitly summon the explicit belief that people do not walk through walls); indeed, surprise cases are a good heuristic precisely in the situation in which one has no occurrent belief having the content of the violated expectation. The puzzle about surprise construed as the result of a mismatch concerns the necessity to postulate that many volatile representations are active at any given moment. If not that many representations are active in this way, then how can we be surprised? But as we can be surprised by so many different things, how many different representations should we postulate as active? The puzzle can be attenuated by arguing that only a limited number of representations are active, precisely ideological representations, whose content is very general. This is corroborated by the observation that we are, in general, much less surprised than a simple mismatch model would imply; it is then plausible to assume that volatile expectations with ideological content are always re-generated and maintained.

## Acknowledgements

Thanks to an anonymous referee for comments on an earlier draft, and to Milena Nuti for countless improvements.

Research for this paper was partly funded through the Enactive Network of Excellence IST-2002-002114.

## References

- Baillargeon, R. (1994). How do infants learn about the physical world. *Current Directions in Psychological Science*, 3(5), 133-140.
- Bain, A. (1874). *The emotions and the will*. New York: Appleton.
- Bruner, J., Postman, L. (1949). On the perception of incongruity: A paradigm. *Journal of Personality*, 18, 206-223.
- Cavanagh, P. (2001). Pictorial art and vision, *MIT Encyclopedia of Cognitive Science* (pp. 644-646). Cambridge, Mass: MIT Press.
- Charlesworth, W. R. (1969). The role of surprise in Cognitive Development. In D. Elkind & J. H. Flavell (Eds.), *Studies in Cognitive Development, essays in honor of Jean Piaget* (pp. 257-314). New-York and London: Oxford University Press.

Andrea Formica 4/12/06 14:11

Supprimé: 11/27/2006

- Coughlan, R., & Connolly, T. (2001). Predicting Affective Responses to Unexpected Outcomes. *Organ Behav Hum Decis Process.*, 85(2), 211-225.
- Craske, B. (1977). Perception of impossible limb position induced by tendon vibration. *Science*, 196, 71-73.
- Darwin, C. (1872). *The expression of emotions in man and animals*. London: Murray.
- Davidson, D. (1982). Rational animals. *Dialectica*, 36, 318-327.
- Davidson, D. (2004). *Problems of rationality*. Oxford: Oxford University Press.
- Dean, G. (2000). *Step by step to stand-up comedy*.: Heinemann Drama.
- Dennett, D. C. (2001). Surprise, surprise," commentary on O'Regan and Noe. *Behavioral and Brain Sciences*, 24(5), 982.
- Gendolla, G. H. E. (1997). Surprise in the context of Achievement: the role of outcome valence and importance. *Cognition and Emotion*, 165-193.
- Heller, M. A., Calcaterra, J. A., Green, S. L., & Brown, L. (1999). Intersensory conflict between vision and touch: the response modality dominates when precise, attention-riveting judgements are required. *Perception & psychophysics*, 61(7), 1384-1398.
- Izard, C. E. (1977). *Human Emotions*. New-York: Plenum Press.
- Izard, C. E. (1991). *The Psychology of Emotions*. New York: Plenum Press.
- Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, 24, 175-219.
- Mc Dougall, W. (1923). *Outline of Psychology*. New York: Scribner's.
- McDaniel, M. A., Einstein, G. O., DeLosh, E. L., May, C. P., & Brady, P. (1995). The bizarreness effect: it's not surprising, it's complex. *J Exp Psychol Learn Mem Cogn.*, 21(2), 422-435.
- Meyer, W.-U., & Niepel, M. (1994). Surprise. In V. S. Ramachandran (Ed.), *Encyclopedia of human behavior* (Vol. 3, pp. 353-358). San Diego: Academic Press.
- Meyer, W.-U., Niepel, M., Rudolph, U., & Schützwohl, A. (1991). An experimental analysis of surprise. *Cognition and Emotion*, 5, 295-311.
- Meyer, W.-U., Reisenzein, R., & Schützwohl, A. (1997). Toward a Process Analysis of Emotions: The case of surprise. *Motivation and Emotion*, 21(3), 251-274.
- Niepel, M., Rudolph, U., & Schützwohl, A. (1994). Temporal characteristics of the surprise reaction induced by schema discrepant visual and auditory events. *Cognition and Emotion*, 8, 433-452.
- Ortony, A., Clore, G. L., & Collins. (1988). *The Cognitive Structure of Emotions*. Cambridge, MA: Cambridge University Press.
- Patterson, E. (1930). A qualitative and quantitative study of the emotion of surprise. *Psychological Monographs*, 40, 85-108.
- Plutchik, R. (1980). *Emotion: Theory, Research and Experience, vol 4, The measurement of emotions*. London: Academic Press.
- Reisenzein, R., W-U., M., & Schützwohl, A. (1996). *Reactions to surprising events: a paradigm for emotion research*. Paper presented at the 9th conference of the International Society for Research on Emotions, Toronto.

Andrea Formica 4/12/06 14:11

Supprimé: 11/27/2006



- Ribot, T. H. (1911). *The psychology of emotions*. London: W.C. Scott, Ltd.
- Rock, I., & Victor, J. (1964). Vision and touch; an experimentally created conflict between the senses. *Science*, *143*, 594-596.
- Ruckmick, C. A. (1921). A preliminary study of the emotions. *Psychological Monographs*, *30*, 30-35.
- Schützwohl, A. (1998). Surprise and Schema Strength. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *24*(5), 1182-1199.
- Spelke, E. S. (1990). Principles of object segregation. *Cognitive Science*, *14*, 29-56.
- Spelke, E. S. (1991). Principles of object perception. *Cognitive Science*, *14*, 29-56.
- Stiensmeier-Pelster, J., Martini, A., & Reisenzein, R. (1995). The role of surprise in the attribution process. *Cognition and Emotion*, *9*, 5-31.
- Sumitsuji, N. (2000). The origin of intermittent exhalation (A! Ha! Ha!) peculiar to human laugh. *Electromyogr Clin Neurophysiol.*, *40*(5), 305-309.
- Vanhamme, J. (1999). La surprise et son influence sur la satisfaction des consommateurs. *Actes du 15e Congrès de l' Association Française du Marketing*, *15*(2), 809-839.
- Vanhamme, J. (2000). The link between surprise and satisfaction: an exploratory research on how best to measure surprise. *Journal of Marketing Management*, *16*, 565-582.
- Warren, H. C. (1920). *Human Psychology*. Boston: Houghton Mifflin.
- Whittlesea, B. W., & Williams, L. D. (2001). The discrepancy-attribution hypothesis: II. Expectation, uncertainty, surprise, and feelings of familiarity. *J Exp Psychol Learn Mem Cogn.*, *27*(1), 14-33.
- Woodworth, R. W., & Schlosberg, H. (1954). *Experimental psychology*. New York: Holt.

Andrea Formica 4/12/06 14:11

Supprimé: 11/27/2006

| 14 Casati Pasquinelli Surprise [12/4/2006](#)